Crystal Structure of the Envelope Glycoprotein GP2 from Ebola Virus at 1.9 Å Resolution

V.N. Malashkevich, B.J. Schneider, M.L. McNally,
M.A. Milhollen, J.X. Pang and P.S. Kim
Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology, Nine Cambridge Center, Cambridge, MA 02142, USA

INTRODUCTION

Ebola viruses are nonsegmented, negative-strand RNA viruses which, together with Marburg viruses, constitute the family *Filoviridae*. The most pathogenic subtype (Zaire) of Ebola virus causes a severe form of hemorrhagic fever in humans and nonhuman primates. Ebola virions possess a single surface transmembrane glycoprotein (GP) that plays a central role in entry into target cells by binding to as yet unidentified receptors, and then mediating fusion between the viral and the host cell membranes. GP is expressed as a single chain precursor of 676 amino acids that is postranslationally processed into the disulfide-linked fragments GP1 and GP2. The first subunit is responsible for binding to cell-surface receptors, while the second subunit is responsible for mediating membrane fusion.

We produced a 95 amino-acid residue fragment of GP2, corresponding to the ectodomain without the amino-terminal fusion-peptide region[1]. After establishing the correct disulfide connectivity in the region of the fragment that corresponds to the retroviral immunosuppressive motif (by characterizing disulfide variants of the ectodomain), limited proteolysis experiments were used to identify a proteolytically resistant core of the ectodomain. This 74 amino acid core (referred to as Ebo-74) was crystallized and the X-ray crystal structure was determined at 1.9 Å resolution[1]. While our crystal structure was being refined, the structure of a similar fragment of Ebola GP2, fused to a trimeric coiled coil, was reported at 3.0 Å resolution[2]. Our results are in overall agreement with the structure of this fusion protein, and they confirm the remarkable structural similarity between the core of the Ebola GP2 protein and the X-ray crystal structure of the corresponding core from the Mo-MLV TM protein (Mo-55) that had been anticipated earlier[3].

MATERIALS AND METHODS

Well-diffracting crystals of Ebo-74 were grown as described[1]. Crystals grew as hexagonal bipyramids to a maximum size of 0.1 mm x 0.1 mm x 0.2 mm within 1 - 2 weeks. The crystals belong to the space group $P6_2$ (or $P6_4$, a= b= 75.66Å, c=67.94Å, $\alpha = \beta = 90^o$, $\gamma = 120^o$) and contain one trimer in the asymmetric unit, wherein monomers are related by an approximate non-crystallographic three-fold axis[1].

For data collection, crystals were transferred to a drop containing reservoir solution plus 30-40% glycerol, harvested, and flash frozen. Diffraction data were collected at 100 K using a Quantum-4 CCD detector and the 5.0.2 beam line at ALS. The Ebo-74 crystal diffraction limits were improved and the mosaicity was decreased significantly by a cryo-annealing technique in which the crystal was repeatedly thawed for about 30 sec by blocking the cryostream and then frozen again. An elevated concentration of cryoprotectant to prevent ice formation was essential with this technique. Diffraction intensities were integrated using the DENZO and SCALEPACK software[4], and reduced to structure factors with the program TRUNCATE from the CCP4 program suite[5]. The structure of Ebo-74 was solved by molecular replacement using the program AMoRe[6]. Solvent flattening, histogram matching and three-fold non-crystallographic averaging with the program DM[7] improved phasing. Crystallographic refinement of the structure (Table 1) was done

Table 1. Data collection and refinement statistics[1]

Data collection	
Resolution, Å	20.0 - 1.90
Observed reflections	86,630
Unique reflections	17,326
Completeness, %	98.9 (99.1)[a]
R_{merge}[b]	0.037 (0.309)[a]
Refinement	
Protein nonhydrogen atoms	1806
Water molecules	214
Heteroatoms (Cl^-)	1
R_{cryst}[c]	0.205 (0.239)[a]
R_{free}[c]	0.269 (0.319)[a]
RMSD from ideal geometry	
Bond lengths, Å	0.009
Bond angles, °	1.3
Dihedral angles, °	19.8
Average B-factor, $Å^2$	40.6

[a]Values in parentheses correspond to highest resolution shell 1.93 to 1.90 Å.

[b]$R_{merge} = \Sigma\Sigma_j|I_j(hkl) - \langle I(hkl)\rangle|/\Sigma|\langle I(hkl)\rangle|$, where I_j is the intensity measurement for reflection j and <I> is the mean intensity over j reflections.

[c]R_{cryst} (R_{free}) = $\Sigma||F_{obs}(hkl)| - |F_{calc}(hkl)||/\Sigma|F_{obs}(hkl)|$, where F_{obs} and F_{calc} are observed and calculated structure factors, respectively. No σ-cutoff was applied. 10 % of the reflections were excluded from refinement and used to calculate R_{free}.

with the CNS[8] and REFMAC[5] programs. Non-crystallographic symmetry restraints were not used in the final refinement[1].

RESULTS

Although Ebo-74 contains significantly more residues than Mo-55[3], and there is only 22% sequence identity between Ebo-74 and Mo-55, the overall similarities of these cores prompted us to attempt to solve the Ebo-74 crystal structure with a molecular replacement approach. A solution was found[1] using a trimeric polyserine model of Mo-55, with the loop regions omitted and a chloride ion retained in the core (i.e., the model contained only 36% of the atoms in the target structure). This solution became apparent after the fitting procedure in AMoRe (final correlation coefficient = 0.431 and R-factor = 0.484 in the resolution shell 10.0 to 4.0 Å), and it also resolved the ambiguity in space group determination in favor of $P6_2$. Initial phases, calculated from the correctly positioned model, were dramatically improved using three-fold averaging with DM. The initial DM-phased electron density map clearly revealed the structures of the omitted loops and of the C-terminal helices absent from the trial model. The quality of the map was further improved through noncrystallographic symmetry matrix refinement and solvent mask redefinition with the program NCSMASK[5]. An initial model of Ebo-74 was built based on this map and then refined to convergence[1].

As expected, Ebo-74 forms a homotrimer. Each of the three polypeptide chains folds into a helical-hairpin conformation, in which two antiparallel helices are connected by a loop region (Fig. 1). The N-terminal helices from each monomer form a central, three-stranded coiled coil. This coiled-coil core includes approximately 35 residues from each chain (the two most N-terminal residues are not well defined in the electron density maps and next two are in a random-coil conformation). Shorter C-terminal helices pack in an antiparallel manner into hydrophobic grooves on the surface of the coiled-coil core. In the loop region connecting the N- and C-terminal helices, a disulfide between Cys-601 and Cys-608 links a short α-helix and a short 3_{10}-helix. Additionally, the loop region between the 3_{10}-helix and C-helix are in an extended conformation. Due to asymmetric crystal contacts, the three individual chains in the Ebo-74 structure have slight differences in conformation and degree of order, with the B chain having a higher overall B-factor (43.1 $Å^2$) than the A and C chains (37.7 and 37.1 $Å^2$, respectively). The RMSD between the individual chains in the Ebo-74 noncrystallographic trimer vary from 0.74 Å (chains A and C) to 1.31 Å (chains B and C)[1].

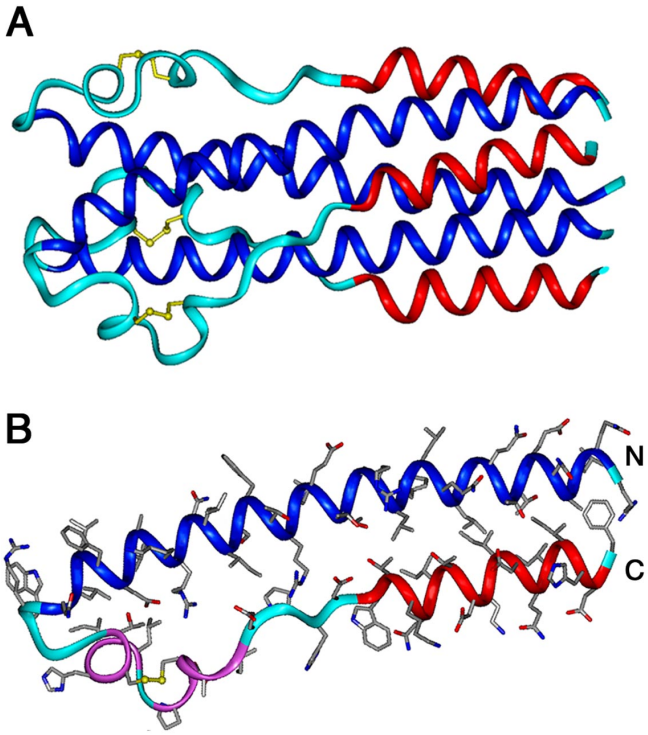

Figure 1. Ebo-74 forms a trimer-of-hairpins structure[1]. (**A**) A side view of the Ebo-74 structure. N-helices (dark blue) constitute a central coiled coil and C-helices (red) pack into hydrophobic grooves on the surface of the coiled coil. Disulfides within the loop regions (light blue) are depicted in yellow. (**B**) Each monomer has an α-helical hairpin conformation.

In the Ebo-74 structure, a strong X-ray scatterer binds between the adjacent rings of Ser-583 and Asn-586 residues. This density was identified as a chloride ion[1]. The individual B-factor refinement of a modeled chloride ion converges at 24.1 Å^2; the same value is obtained for the average B-factor of the interacting amide nitrogens of the Asn-586 residues. The average distance between the ion and the interacting nitrogen atoms is 3.28 Å, similar to the distances observed for chloride binding in other coiled coils and close to the sum of the van der Waals radii for NH_4^+ and Cl^- (3.24 Å). If a chloride-binding site is common in other fusion glycoproteins with analogous architectures and can be substituted by bromide, the bromide complex would be an attractive target for MAD phasing experiments when molecular replacement methods fail to permit X-ray structure determination.

ACKNOWLEDGMENTS

We thank Dr. Thomas Earnest for beamtime allocation and help at the 5.0.2 beamline.

REFERENCES

1. V.N. Malashkevich *et al.* Proc. Natl. Acad. Sci. USA **96**, in press (1999).
2. W. Weissenhorn *et al.*, Molec. Cell **2**, 605-616 (1998).
3. D. Fass, S.C. Harrison, and P.S. Kim, Nat. Struct. Biol. **3,** 465-469 (1996) .
4. Otwinowski, Z., in *Data Collection and Processing*, eds. Sawer, L., Isaacs, N. & Bailey, S. (SERC, Daresbury Laboratory, Warrington, England, 1993), pp. 55-62.
5. CCP4 Acta Crystallogr. **D50**, 760-763 (1994).
6. J. Navaza, Acta Crystallogr. **A50,** 157-163 (1994).
7. K.D. Cowtan, Joint CCP4 and ESF-EACBM Newslett. Prot. Cryst. **31**, 43-38 (1994).
8. A.T. Brünger, *et al.* Acta Crystallogr. **D54,** 905-921 (1998).

This research was funded by the NIH (PO1 GM56552).

Principal investigator: Peter S. Kim, Howard Hughes Medical Institute, Whitehead Institute for Biomedical Research, Department of Biology, Massachusetts Institute of Technology. Email: tocio@wi.mit.edu. Telephone: 617-258-5184.